Lego Case Study - Random Effects Practice

STAT 245

Data

The dataset for today gives details on `Price` and other characteristics of 6172 Lego sets, including the number of `Pieces` and `Minifigures` inside and the `Year` they were created, as well as the `Name`, `Theme`, `Subtheme` and `Item_Number` of each set.

Find the data at: https://sldr.netlify.app/data/legos.csv.

Figure 1: Some Lego bricks and minifigures. Image is from https://www.ikea.com/us/en/p/bygglek-201-piece-lego-r-brick-set-mixed-colors-30455758/

Tasks

Do the jobs below, saving your work in a Quarto document.

1. **Fit a regression model to this dataset.** You can choose the response variable. I suggest choosing from: `price`, `Pieces`, or `Minifigures`. Do at least a short version of model planning, but don't spend much of your time on planning or graphs of the data (prioritize time for fitting and prediction).
2. **Include at least one random effect in your model.**
3. **Short model assessment** (*Focusing on ACF only because today's main goals are to practice model* fitting *and* predictions).

- Fit two versions of your model:
  - One *with* the random effect (as planned)
  - One with all random effect variables *completely excluded*
- Show ACFs for both of the models. Does adding the random effect(s) help reduce residual dependence?

4. **Choose** whether you want to make *average random effect group* or *population average* predictions. **Explain why in words, in context** (that is, using words about Legos to illustrate what the "typical RE group" or the "population average" *mean*).
5. **Make a prediction plot** that makes the kind of predictions you chose. (This will either be using `ggpredict()` as usual, or doing a parametric bootstrap.)